

Supratik Padhee

+91 7978259993 | supratikpadhee.career@outlook.com

[LinkedIn](#) | [GitHub](#)

PROFILE

Skilled and Result-Driven, Data and Platform Engineer with 8+ years of experience in designing and optimizing pipelines, ETL workflows, and enterprise data warehouse solutions. Proficient in Python, SQL, Flask, FastApi, pySpark, Kubernetes, Docker, Databricks, Kafka, AWS EMR, and AWS Glue, AWS RDS, Snowflake to build scalable, multi-terabyte big data platforms for companies in Supply-Chain, Ecommerce, Entertainment, Transportaion domains.

TECHNICAL SKILLS

Programming Languages:	Python, SQL, Scala
Big Data Technologies:	Spark, PySpark, Spark SQL, YARN, Kubernetes, Hadoop, Hive, Delta Lake
Streaming Technology:	Kafka, Spark structured streaming
Cloud computing:	Databricks, Amazon S3, AWS Glue, AWS EMR, AWS Lambda, AWS Dynamodb
Data Engineering Tools:	Data Modelling, ETL/ELT data Pipeline, Apache Zepellin
Orchestration:	Airflow, Kubernetes
Familiar:	FastAPI, Flask Api, Elastic search, Neo4J

EXPERIENCE

Persistent Systems

Senior Data Engineer - Various Projects

May 2025 - Present

Bengaluru, India

- **Unified Ingestion Platform**

- * Developed, optimized **ETL pipelines** using **spark, Hive, PySpark (DataFrame, S3 SparkSQL, RDD)**, to batch replicate data from RDS/NoSQL data sources to centralised datalake.

Brambles

Senior Data Engineer - Various Projects

December 2021 - May 2025

Bengaluru, India

- **InsightsToAction Analysis Platform**

- * Developed, optimized **ETL pipelines** using **PostgreSQL, Hive, PySpark (DataFrame, S3 SparkSQL, RDD)**, orchestrated with **Apache Airflow**, and enabled insights through **Apache Superset** visualizations.

- **Job Processing and Asset Tracking Core**

- * Developed **AWS ECS Flask API** server generating **dynamic DAGs** in **Apache Airflow**, push messages into a **Kafka** broker for real-time tracking of DAG processing status (success or failure).

AltraDimension Technologies Private Limited

September 2020 - December 2021

Cloud Data Platform Engineer

Remote, India

- **Identity Resolution**

- * Engineered Identity Resolution system on **AWS (S3, Glue, Lambda, Neptune, CloudFormation, DynamoDB, Elasticsearch, API Gateway)** using incremental, historical using **Kafka**, and **S3**, creating a unified customer identity.

Xerox Technology Services

September 2017 - May 2020

Data Platform Engineer

Bengaluru, India

- **CiteWeb and RedLight Analysis**

- * Engineered scalable computing platform using **Python, Flask, AWS EMR, Spark, Zeppelin, Apache Hive, Livy, and Ranger**, enabling users to define, execute, and manage data analytics jobs, improving efficiency.

EDUCATION

CDAC Electronics-city, Bengaluru

Bengaluru, Odisha

Post Graduate Diploma in Advance Computing (PGDAC)

January 2017 - September 2017

Silicon University, Bhubaneshwar

Bhubaneshwar, Odisha

Bachelor of Technology in Computer Science and Engineering (BTECH)

September 2011 - May 2015